(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 September 2001 (13.09.2001)

PCT

(10) International Publication Number
WO 01/66693 A1

(51) International Patent Classification[7]: C12N 1/19, 15/81

(21) International Application Number: PCT/DK01/00154

(22) International Filing Date: 9 March 2001 (09.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

PA 2000 00392	10 March 2000 (10.03.2000)	DK
PA 2000 00419	15 March 2000 (15.03.2000)	DK

(71) Applicant: **NOVOZYMES A/S** [DK/DK]; Patents, Krogshøjvej 36, DK-2880 Bagsværd (DK).

(72) Inventors: **TAKAGI, Shinobu**; No. 708, 1-31-1, Maehara-Nishi, Funabashi-shi, Chiba 274-0825 (JP). **TERUI, Yuji**; 1-46-12 Chishirodai-nishi, Wakaba Ward, Chiba, Chiba 264-0004 (JP). **TSUTSUMI, Noriko**; 3-2-16, Higashisugano, Ichikawa-shi, Chiba-ken, Chiba 272 (JP). **TAIRA, Rikako**; 9-10, Taihei 4-chome, Sumida-ku, Tokyo 130-0012 (JP).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*
— *with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/66693 A1

(54) Title: COMPOSITIONS AND METHODS FOR PRODUCING HIGH YIELDS OF HETEROLOGOUS POLYPEPTIDES IN A *PICHIA* CELL

(57) Abstract: The present invention relates to a method of heterologous production in Pichia, where the codon usage of a heterologous nucleotide sequence is adjusted to the preferred codon usage of Pichia methanolica.

Compositions and methods for producing high yields of heterologous polypeptides in a *Pichia* cell

Field of invention

The high yield production of certain polypeptides is of industrial importance, thus compositions and methods for producing high yields of heterologous polypeptides in a *Pichia* cell such as *Pichia methanolica* are of interest.

Background of the invention

Yeast cells are potential candidates for host organisms for heterologous protein production. To use *Saccharomyces cerevisiae* for the production of e.g. pharmaceutical products or industrial enzymes is well known. It has been suggested that methylotrophic yeasts that can utilize methanol as a sole carbon and energy source are more attractive host strains for heterologous protein production. High protein productivity has been achieved with strongly inducible promoters, such as the one from alcohol oxidase or formate dehydrogenase, by performing high cell density fermentations in which the dried cell mass exceeds 100 g/kg. Examples of high expression of heterologous protein are: 2.5 g/l of *S.cerevisiae* invertase (Tschopp et. al. Bio/Technology, vol 5, 1305-1308, 1987); 3.4 g/l of human serum albumin in *Pichia pastoris* (EP 510693); 3.4g/l of glucoamylase from *Rhizopus oryzae* in *Candida boidinii* (Sakai et.al, Biochem. Biophys. Acta, 1308, 81-87,1996); 1.4g/l of *Schawanniomyces occidentalis* glucoamylase (Gelissen et.al, Bio/Technology, vol 9, 291-295, 1991); and 7.2g/l of *Aspergillus* phytase in *Hansenula polymorpha* (EP911416). In WO 99/67398 an *Aspergillus niger* phytase is expressed in and secreted from *Pichia pastoris* by using a fusion with the signal peptide of yeast α-factor.

Pichia methanolica is another methylotrophic yeast that has been shown

recently to produce human glutamate decarboxylase at the level of 0.5 g/l intracellularly (Yeast 14: 11-23. 1998). It has been stated that *P. methanolica* could be used for production of any heterologous protein, but no example of productivity or yield over 1 gram per litre has been shown yet (US 5,716,808; WO 97/17450).

Summary of the invention

The problem to be solved by this invention is to provide expression systems and methods for producing a polypeptide of interest in high yields.

The solution is based on that the present inventors have identified compositions and methods for high yield production of a heterologous polypeptide of interest in a *Pichia* cell specifically *Pichia methanolica*. By adjusting the codon usage of a gene encoding a polypeptide of interest to the codon usage of one or more gene(s) from the wild type *P methanolica*, the inventors have shown that protein productivity surprisingly can be increased with as much as a factor of 10 or more.

In non-limiting example herein heterologous polypeptides were expressed from codon-optimized genes with original secretion signal sequences and secreted from *P. methanolica*.

In the industrial production of polypeptides, a gene encoding a polypeptide of interest is most often present within the production host cell in multiple copies which are all expressed in order to achieve an optimally high yield of the polypeptide. Gene analysis by Southern blotting of a production cell of the invention revealed that the high polypeptide yield was achieved with a very low gene dosage of only approx. 2 copies pr. cell, which suggests a potential for increasing the yield of the polypeptide of interest even further by increasing the copy number.

Accordingly in a first aspect, the present invention relates to a *Pichia* cell comprising at least one copy of a heterologous nucleotide sequence encoding a

polypeptide of interest, wherein the codon usage of said sequence has been adjusted to match the preferred codon usage of *P. methanolica*, as defined herein.

The preferred codon usage of *P. methanolica* as relates to this invention is presented in Example 2, Table 1, e.g. as the codon-ratios used in the optimized phytase gene sequence for each amino acid. More specifically the preferred codons for Lysine are TTA or TTG; for Isoleucine they are ATT or ATC; for Valine they are GTT or less preferred GTC; for Serine they are TCT, TCC or TCA; for Proline they are CCA or less preferred CCT; for Threonine they are ACT or less preferred ACC; for Alanine they are GCT or GCC; for Tyrosine the more preferred is TAC; for Glutamine it is CAA; for Asparagine it is AAC; for Glutamic acid it is GAA; for Cysteine it is TGT; for Arginine it is AGA; and finally for Glycine it is GGT.

The codon usage of a DNA sequence encoding a polypeptide is adjusted to match the preferred codon usage of *P. methanolica* when at least 50% of the codons for specific amino acids in the polypeptide sequence have been adjusted to a preferred codon as defined above, or preferably at least 60%, more preferably at least 70%, even more preferably 80%, yet more preferably 90%, and most preferably 95% of the codons for specific amino acids in the polypeptide sequence have been adjusted to a preferred codon as defined above.

Further in a second aspect, the present invention relates to an isolated DNA construct comprising at least one copy of a nucleotide sequence encoding a polypeptide heterologous to *Pichia methanolica*, wherein the codon usage of said sequence has been adjusted to match the preferred codon usage of *P. methanolica*, as defined herein.

Finally in a third aspect the present invention relates to a method of producing a polypeptide of interest in a *Pichia* cell, where the polypeptide is encoded by a nucleotide sequence heterologous to *P. methanolica*, said method comprising the steps of:

a) adjusting the codon usage of the sequence to match the preferred codon usage of *P.*

methanolica, as defined herein; and

b) cultivating a *Pichia* cell comprising at least one copy of the codon usage adjusted sequence of step a) under appropriate growth conditions to express the sequence and achieve production of the polypeptide of interest.

Deposited microorganisms

The entire sequence of the synthetic codon optimized *A. fumigatus* phytase gene (SEQ ID No. 19) is comprised in the strain *E. coli* NN049526, which was deposited with "Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" (DSMZ) 01-March-2000 with the accession number DSM 13352.

The cDNA sequence of the *Thermomyces lanuginosus* DSM4109 xylanase gene was published in WO 96/32472. For this study we obtained a *Saccharomyces cerevisiae* strain from DSMZ, accession number DSM 10133, deposited on 19-July-1995, which contains a plasmid comprising the full-length xylanase cDNA of the *Thermomyces lanuginosus* DSM4109.

The deposits were made by Novo Nordisk A/S and were later assigned to Novozymes A/S.

Definitions

The term "isolated nucleic acid sequence" as used herein refers to a nucleic acid sequence which is essentially free of other nucleic acid sequences, *e.g.*, at least about 20% pure, preferably at least about 40% pure, more preferably at least about 60% pure, even more preferably at least about 80% pure, and most preferably at least about 90% pure as determined by agarose electrophoresis.

For example, an isolated nucleic acid sequence can be obtained by standard cloning procedures used in genetic engineering to relocate the nucleic acid sequence

from its natural location to a different site where it will be reproduced. The cloning procedures may involve excision and isolation of a desired nucleic acid fragment comprising the nucleic acid sequence encoding the polypeptide, insertion of the fragment into a vector molecule, and incorporation of the recombinant vector into a host cell where multiple copies or clones of the nucleic acid sequence will be replicated.

The nucleic acid sequences of the present invention may be of genomic, cDNA, RNA, semisynthetic, synthetic origin, or any combinations thereof and may be obtained from microorganisms of any genus. For purposes of the present invention, the term "obtained from" as used herein in connection with a given source shall mean that the polypeptide encoded by the nucleic acid sequence is produced by the source or by a cell in which the nucleic acid sequence from the source has been inserted.

The nucleic acid sequences may be obtained from a bacterial source. For example, these polypeptides may be obtained from a gram positive bacterium such as a *Bacillus* strain, *e.g.*, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis*, or *Bacillus thuringiensis*; or a *Streptomyces* strain, *e.g.*, *Streptomyces lividans* or *Streptomyces murinus*; or from a gram negative bacterium, *e.g.*, *E. coli* or *Pseudomonas* sp.

The nucleic acid sequences may be obtained from a fungal source, and more preferably from a yeast strain such as a *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, or *Yarrowia* strain; or more preferably from a filamentous fungal strain such as an *Acremonium*, *Aspergillus*, *Aureobasidium*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Piromyces*,

Schizophyllum, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, or *Trichoderma* strain.

The nucleic acid sequences may thus be obtained from *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* or *Saccharomyces oviformis*; or indeed from from an *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, or *Trichoderma viride* strain.

In a preferred embodiment, the nucleic acid sequence is the sequence encoding a phytase comprised in *Escherichia coli* DSM13352. In another preferred embodiment, the nucleic acid sequence is SEQ ID NO: 3 or SEQ ID NO: 19.

It will be understood that for the aforementioned species, the invention encompasses both the perfect and imperfect states, and other taxonomic equivalents, *e.g.*, anamorphs, regardless of the species name by which they are known. Those skilled in the art will readily recognize the identity of appropriate equivalents. Strains of these species are readily accessible to the public in a number of culture collections, such as the American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Centraalbureau Voor

Schimmelcultures (CBS), and Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

Furthermore, such nucleic acid sequences may be identified and obtained from other sources including microorganisms isolated from nature (*e.g.*, soil, composts, water, etc.) using the above-mentioned probes. Techniques for isolating microorganisms from natural habitats are well known in the art. The nucleic acid sequence may then be derived by similarly screening a genomic or cDNA library of another microorganism. Once a nucleic acid sequence encoding a polypeptide has been detected with the probe(s), the sequence may be isolated or cloned by utilizing techniques which are known to those of ordinary skill in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

The techniques used to isolate or clone a nucleic acid sequence encoding a polypeptide are known in the art and include isolation from genomic DNA, preparation from cDNA, or a combination thereof. The cloning of the nucleic acid sequences of the present invention from such genomic DNA can be effected, *e.g.*, by using the well-known polymerase chain reaction (PCR) or antibody screening of expression libraries to detect cloned DNA fragments with shared structural features. See, *e.g.*, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Other nucleic acid amplification procedures such as ligase chain reaction (LCR), ligated activated transcription (LAT) and nucleic acid sequence-based amplification (NASBA) may be used.

A nucleic acid sequence of the present invention may also encode fused polypeptides or cleavable fusion polypeptides in which another polypeptide is fused at the N-terminus or the C-terminus of the polypeptide or fragment thereof. A fused polypeptide is produced by fusing a nucleic acid sequence (or a portion thereof) encoding another polypeptide to a nucleic acid sequence (or a portion thereof) of the

present invention. Techniques for producing fusion polypeptides are known in the art, and include ligating the coding sequences encoding the polypeptides so that they are in frame and that expression of the fused polypeptide is under control of the same promoter(s) and terminator.

The present invention also relates to nucleic acid constructs comprising a nucleic acid sequence of the present invention operably linked to one or more control sequences which direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences. Expression will be understood to include any step involved in the production of the polypeptide including, but not limited to, transcription, post-transcriptional modification, translation, post-translational modification, and secretion.

"Nucleic acid construct" or "DNA construct" is defined herein as a nucleic acid molecule, either single- or double-stranded, which is isolated from a naturally occurring gene or which has been modified to contain segments of nucleic acid which are combined and juxtaposed in a manner which would not otherwise exist in nature. The term nucleic acid construct is synonymous with the term expression cassette when the nucleic acid construct contains all the control sequences required for expression of a coding sequence of the present invention. The term "coding sequence" is defined herein as a portion of a nucleic acid sequence which directly specifies the amino acid sequence of its protein product. The boundaries of the coding sequence are generally determined by a ribosome binding site (prokaryotes) or by the ATG start codon (eukaryotes) located just upstream of the open reading frame at the 5' end of the mRNA and a transcription terminator sequence located just downstream of the open reading frame at the 3' end of the mRNA. A coding sequence can include, but is not limited to, DNA, cDNA, and recombinant nucleic acid sequences.

An isolated nucleic acid sequence encoding a polypeptide of the present invention may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the nucleic acid sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying nucleic acid sequences utilizing recombinant DNA methods are well known in the art.

The term "control sequences" is defined herein to include all components which are necessary or advantageous for the expression of a polypeptide of the present invention. Each control sequence may be native or foreign to the nucleic acid sequence encoding the polypeptide. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator. At a minimum, the control sequences include a promoter, and transcriptional and translational stop signals. The control sequences may be provided with linkers for the purpose of introducing specific restriction sites facilitating ligation of the control sequences with the coding region of the nucleic acid sequence encoding a polypeptide. The term "operably linked" is defined herein as a configuration in which a control sequence is appropriately placed at a position relative to the coding sequence of the DNA sequence such that the control sequence directs the expression of a polypeptide.

The control sequence may be an appropriate promoter sequence, a nucleic acid sequence which is recognized by a host cell for expression of the nucleic acid sequence. The promoter sequence contains transcriptional control sequences which mediate the expression of the polypeptide. The promoter may be any nucleic acid sequence which shows transcriptional activity in the host cell of choice including mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the

host cell.

Useful signal peptides for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding regions are described by Romanos *et al.*, 1992, *Yeast* 8: 423-488.

It may also be desirable to add regulatory sequences which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound.

The present invention also relates to recombinant expression vectors comprising a nucleic acid sequence of the present invention, a promoter, and transcriptional and translational stop signals. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present invention may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

The recombinant expression vector may be any vector (*e.g.*, a plasmid or virus) which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the nucleic acid sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids.

The vector may be an autonomously replicating vector, *i.e.*, a vector which exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.*, a plasmid, an extrachromosomal element, a mini-chromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

The vectors of the present invention preferably contain one or more selectable markers which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs, and the like.

Suitable markers for yeast host cells are ADE1, ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3.

The vectors of the present invention preferably contain an element(s) that permits stable integration of the vector into the host cell genome or autonomous replication of the vector in the cell independent of the genome of the cell. For integration into the host cell genome, the vector may rely on the nucleic acid sequence encoding the polypeptide or any other element of the vector for stable integration of the vector into the genome by homologous or nonhomologous recombination. Alternatively, the vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). To increase the likelihood of integration at a precise location, the integrational elements should preferably contain a

sufficient number of nucleic acids, such as 100 to 1,500 base pairs, preferably 400 to 1,500 base pairs, and most preferably 800 to 1,500 base pairs, which are highly homologous with the corresponding target sequence to enhance the probability of homologous recombination. The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences. On the other hand, the vector may be integrated into the genome of the host cell by non-homologous recombination.

For autonomous replication, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of origins of replication for use in a yeast host cell are the two micron origins of replication, ARS1, ARS4, the combination of ARS1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes its functioning temperature-sensitive in the host cell (see, *e.g.*, Ehrlich, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1433).

More than one copy of a nucleic acid sequence of the present invention may be inserted into the host cell to increase production of the gene product. An increase in the copy number of the nucleic acid sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

The procedures used to ligate the elements described above to construct the recombinant expression vectors of the present invention are well known to one skilled in the art (see, *e.g.*, Sambrook *et al.*, 1989, *supra*).

The present invention also relates to recombinant host cells, comprising a nucleic acid sequence of the invention, which are advantageously used in the recombinant production of the polypeptides. A vector comprising a nucleic acid sequence of the present invention is introduced into a host cell so that the vector is maintained as a chromosomal integrant or as a self-replicating extra-chromosomal vector as described earlier. The term "host cell" encompasses any progeny of a parent cell that is not identical to the parent cell due to mutations that occur during replication. The choice of a host cell will to a large extent depend upon the gene encoding the polypeptide and its source.

Fungal cells may be transformed by a process involving protoplast formation, transformation of the protoplasts, and regeneration of the cell wall in a manner known *per se*. Suitable procedures for transformation of *Aspergillus* host cells are described in EP 238 023 and Yelton *et al.*, 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Suitable methods for transforming *Fusarium* species are described by Malardier *et al.*, 1989, *Gene* 78: 147-156 and WO 96/00787. Yeast may be transformed using the procedures described by Becker and Guarente, *In* Abelson, J.N. and Simon, M.I., editors, *Guide to Yeast Genetics and Molecular Biology*, *Methods in Enzymology*, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, *Journal of Bacteriology* 153: 163; and Hinnen *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

The present invention also relates to methods for producing a polypeptide comprising cultivating a host cell under conditions suitable for production of the polypeptide. In the production methods of the present invention, the cells are cultivated in a nutrient medium suitable for production of the polypeptide using methods known in the art. For example, the cell may be cultivated by shake flask cultivation, small-scale or large-scale fermentation (including continuous, batch,

fed-batch, or solid state fermentations) in laboratory or industrial fermentors performed in a suitable medium and under conditions allowing the polypeptide to be expressed and or isolated. The cultivation takes place in a suitable nutrient medium comprising carbon and nitrogen sources and inorganic salts, using procedures known in the art. Suitable media are available from commercial suppliers or may be prepared according to published compositions (*e.g.*, in catalogues of the American Type Culture Collection). If the polypeptide is secreted into the nutrient medium, the polypeptide can be recovered directly from the medium. If the polypeptide is not secreted, it can be recovered from cell lysates.

The polypeptides may be detected using methods known in the art that are specific for the polypeptides. These detection methods may include use of specific antibodies, formation of an enzyme product, or disappearance of an enzyme substrate. For example, an enzyme assay may be used to determine the activity of the polypeptide as described herein.

The resulting polypeptide may be recovered by methods known in the art. For example, the polypeptide may be recovered from the nutrient medium by conventional procedures including, but not limited to, centrifugation, filtration, extraction, spray-drying, evaporation, or precipitation.

The polypeptides may be purified by a variety of procedures known in the art including, but not limited to, chromatography (*e.g.*, ion exchange, affinity, hydrophobic, chromatofocusing, and size exclusion), electrophoretic procedures (*e.g.*, preparative isoelectric focusing), differential solubility (*e.g.*, ammonium sulfate precipitation), SDS-PAGE, or extraction (see, *e.g.*, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

Detailed description of the invention

A *Pichia* cell comprising at least one copy of a heterologous nucleotide sequence encoding a polypeptide of interest, wherein the codon usage of said sequence has been adjusted to match the preferred codon usage of *P. methanolica*, as defined herein.

The specific *Pichia* cell may be any *Pichia* cell, such as a *Pichia methanolica*, *Pichia pastoris*, *Pichia anomala*, *Pichia stipitis*, *Pichia angusta* or *Pichia guilliermondii* cell. Preferably the specific *Pichia* cell is a *Pichia methanolica* cell.

A preferred embodiment of the invention relates to a *Pichia* cell of the first aspect of the invention, wherein the nucleotide sequence encodes an enzyme, preferably the nucleotide sequence encodes an enzyme having phytase or xylanase activity, or even more preferably the nucleotide sequence comprises a DNA sequence at least 90% identical, or preferably at least 93% identical, or more preferably at least 95% identical, or most preferably at least 97% identical to the sequences shown in SEQ ID No. 3, SEQ ID No. 19, or to the phytase encoding sequence comprised in a cell of DSM 13352.

As is discussed elsewhere herein, there are a number of ways to introduce a heterologous DNA sequence into a host cell known in the art.

A further embodiment relates to a *Pichia* cell of the first aspect, wherein the nucleotide sequence is present on an extrachromosomal DNA construct, preferably a plasmid, or more preferably the nucleotide sequence is integrated in the cell's chromosome, preferably in multiple copies.

Also well known in the art are ways of expressing heterologous DNA sequences encoding polypeptides of interest in host cells. Several promoters have been described to work well in *Pichia* cells as described above.

Accordingly a preferred embodiment of the invention relates to a *Pichia* cell of the first aspect, wherein the nucleotide sequence is transcribed from a promoter of a

methanol-inducible *P. methanolica* gene, preferably the promoter is a native promoter of a *P. methanolica* gene encoding an enzyme selected from the group consisting of alcohol oxidase, dihydroxyacetone synthase, formate dehydrogenase, and catalase.

For industrial production of polypeptides it is of considerable interest to have the polypeptides secreted from the producing host cells in order to minimize costs of isolating the polypeptide, ways of promoting efficient secretion are well known in the art. A polypeptide may be efficiently secreted if it carries an effective secretion signal peptide recognized by the host cell's secretion apparatus. A number of secretion signal peptides have been described to work well in yeast host cells as described elsewhere herein. The present inventors found that even without the addition of a well-characterized yeast secretion signal peptide, efficient secretion could be achieved from a host cell of a polypeptide carrying its native original secretion signal sequence.

A preferred embodiment of the invention relates to a *Pichia* cell of the first aspect, wherein the polypeptide comprises its native original secretion signal sequence.

The present invention relies on adjusting the codon usage of an isolated DNA sequence to match the preferred codon usage of *Pichia methanolica*. An isolated DNA construct comprising at least one copy of a nucleotide sequence encoding a polypeptide heterologous to *Pichia methanolica*, wherein the codon usage of said sequence has been adjusted to match the preferred codon usage of *P. methanolica*, as defined herein.

A preferred embodiment of the invention relates to an isolated DNA construct of the second aspect of the invention, wherein the nucleotide sequence encodes an enzyme, preferably the nucleotide sequence encodes an enzyme having phytase or xylanase activity, and more preferably the nucleotide sequence comprises a DNA sequence at least 90% identical, or preferably at least 93% identical, or more preferably at least 95% identical, or most preferably at least 97% identical to the sequences shown

in SEQ ID No. 3, SEQ ID No. 19, or to the phytase encoding sequence comprised in a cell of DSM 13352.

Another preferred embodiment of the invention relates to an isolated DNA construct of the second aspect of the invention, wherein the nucleotide sequence is transcribed from a promoter of a methanol-inducible *P. methanolica* gene, preferably the promoter is a native promoter of a *P. methanolica* gene encoding an enzyme selected from the group consisting of alcohol oxidase, dihydroxyacetone synthase, formate dehydrogenase, and catalase. Other promoters generally recognized as working well in *Pichia* cells are those of the glycolytic enzymes, such as glyceraldehyde-3-phosphate dehydrogenase, triose phosphate isomerase, enolase etc.

Further a preferred embodiment of the invention relates to an isolated DNA construct of the second aspect of the invention, wherein the nucleotide sequence encodes a polypeptide comprising its native original secretion signal sequence.

As mentioned previously the invention relates to industrial production, *i.e.* a method of producing a polypeptide of interest in a *Pichia* cell, where the polypeptide is encoded by a nucleotide sequence heterologous to *P. methanolica*, said method comprising the steps of:

a) adjusting the codon usage of the sequence to match the preferred codon usage of *P. methanolica*, as defined herein; and

b) cultivating a *Pichia* cell comprising at least one copy of the codon usage adjusted sequence of step a) under appropriate growth conditions to express the sequence and achieve production of the polypeptide of interest.

A preferred embodiment of the invention relates to a method of the third aspect, wherein the *Pichia* cell is a *Pichia methanolica* cell.

Yet another preferred embodiment of the invention relates to a method of the third aspect, wherein the nucleotide sequence encodes an enzyme, preferably the

nucleotide sequence encodes an enzyme having phytase or xylanase activity, or more preferably the nucleotide sequence comprises a DNA sequence at least 90% identical, or preferably at least 93% identical, or more preferably at least 95% identical, or most preferably at least 97% identical to the sequences shown in SEQ ID No. 3, SEQ ID No. 19, or to the phytase encoding sequence comprised in a cell of DSM 13352.

The present inventors succeeded in obtaining much improved yields of polypeptides of interest by using a method of the third aspect of the invention, in one non-limiting example they even obtained a yield that exceeded 10 times the yield obtained with the conventional method.

Consequently a preferred embodiment relates to a method of the third aspect, wherein the polypeptide is produced in a yield at least three times higher than what is obtained from an otherwise identical cell comprising a sequence not adjusted to the preferred codon usage of *P. methanolica*, as defined herein, preferably at least five times higher, more preferably seven, even more preferably eight, yet more preferably nine, and most preferably ten times higher than what is obtained from an otherwise identical cell comprising a sequence not adjusted to the preferred codon usage of *P. methanolica*, as defined herein.

In parallel with the two first aspects, preferred embodiments relate to methods of the third aspect, wherein the nucleotide sequence is transcribed from a promoter of a methanol-inducible *P. methanolica* gene, preferably the promoter is a native promoter of a *P. methanolica* gene encoding an enzyme selected from the group consisting of alcohol oxidase, dihydroxyacetone synthase, formate dehydrogenase, and catalase; and preferably methanol is added to the cell during the cultivating step, whereby the methanol-inducible promoter is induced; it is also preferred that the nucleotide sequence encodes a polypeptide comprising its native original secretion signal sequence.

Material and methods

Purchased material (*E.coli*, plasmid and kits)

E.coli DH12S (Gibco BRL) and *E.coli* JM109 (Toyobo) were used for plasmid construction and amplification. Amplified plasmid was recovered with Qiagen® Plasmid Kit (Qiagen). Ligation was done with DNA ligation kit (Takara) or T4 DNA ligase (Boehringer Mannheim). Polymerase Chain Reaction (PCR) was carried out either with Taq DNA polymerase (Boehringer Mannheim) or with Expand ™ PCR system especially Expand™ Long Template (Boehringer Mannheim). QIAquick™ Gel Extraction Kit (Qiagen) was used for the purification of PCR fragments and extraction of DNA fragment from agarose gel. A plasmid pT7blue T-vector (Novagen) was used for the cloning of PCR fragments.

Expression host strain

The expression host strain is a *Pichia methanolica* PMAD16 (*ade2- pep4Δprb1Δ*) described in WO 99/14347. It is a vacuolar protease deficient mutant originally generated from *P. methanolica* CBS 6515. Two major vacuolar proteinases, i.e. Proteinase A encoded by the pep4 gene and Proteinase B encoded by the prb1 gene, are inactivated by disrupting these genes (WO 99/14347).

Expression plasmid

The expression cassette plasmid pCZR134 is described in WO 99/14347. It comprises the promoter and terminator of the alcohol oxidase gene AUG1 from *P. methanolica* (WO 99/14347) and the ADE2 gene encoding phosphoribosyl-5-aminoimidazole carboxylase (WO 99/14347) from *P. methanolica*.

Another expression cassette plasmid is pGAP, which is identical to pCZR134 but contains the promoter of the gene encoding the glyceraldehyde-3-phophate

dehydrogenase GAP1 from *P. methanolica* (described in WO00/56903. SEQ ID No.8 therein) instead of the AUG1 promoter as described above.

Transformation of *P. methanolica*

Transformation of *Pichia methanolica* can be done using the general methods for yeast transformation. As the most preferred method, the procedure of electro transformation is described below. The *P. methanolica* host strain is grown in non-selective YPD medium (20g/l of Pepton, Difco, 10g/l of yeast extract, Difco and 20g/l of glucose) at 37°C for overnight with vigorous aeration, and the culture is diluted 100-fold in 200ml fresh, pre-warmed YPD. This is further cultivated at 37°C until the optical density at 663nm becomes 1.2-1.5, which corresponds to $9x10^7$ cells/ml. Cells are collected by centrifugation at 3000 g for 10 minutes, and re-suspended in 40ml of 50mM potassium phosphate buffer pH 7.5 containing 25mM of dithiothreitol (DTT), then incubated at 37°C for 15 minutes. The cells are washed twice with ice-cold electroporation buffer STM (270mM sucrose, 10mM Tris pH7.5 and 1mM CaCl2), at first with 200 ml (the original culture volume), then with 100 ml (50% original volume). Finally the cells are re-suspended in 1 ml (0.5% of original volume) ice-cold STM buffer to a concentration of approximately $2x10^{10}$ electrocompetent cells/ml.

0.1-10 µg of DNA (volume < 10µl) is added to 100µl of electrocompetent cells and placed into an ice-cold electroporation cuvette with an electrode gap of 2mm. A pulsed electric field is provided at 3.75 kV/cm for 20 millisecond, using a BioRad Gene Pulser™ electroporator on which resistance is set to >600 Ω or "infitine" and capacitance is set to 25 µF. After delivering the pulse, 1ml of YPD is added to the cell/DNA mixture and incubated for 1 hour at 30°C.

Cells are harvested by centrifugation at 3000 g for 5 minutes, then washed with 1ml of yeast nitrogen base (6.7 g/l yeast nitrogen base without amino acids, Difco), and

plated on minimal selective media, e.g. SC glucose. Host cells having an ade2 mutation appear as pink colonies on minimal medium, whereas transformants that acquired a stable integration (Ade+) can grow rapidly as white colonies on minimal medium.

Example 1

A cDNA clone of the *Thermomyces* xylanase gene

The cDNA sequence of the *Thermomyces lanuginosus* DSM4109 xylanase gene was published (WO 96/32472). For this study we obtained A *Saccharomyces cerevisiae* strain, DSM 10133, containing a plasmid comprising the full-length xylanase cDNA of the *Thermomyces lanuginosus* DSM4109.

Construction of expression plasmid pCaXI1

The *Thermomyces* xylanase cDNA was amplified by PCR using primers xyN and xyC to introduce desirable restriction enzyme cleavage sites at both ends of the PCR fragment:

xyN (SEQ ID 1): CCGGAATTCATGGTCGGCTTTACCCCCGTT.

XyC (SEQ ID 2): AAGGAAAAAAGCGGCCGCACTAGTTTAGCCCA CGTCAGCAAC.

The PCR reaction mixture comprised 50 ng of template cDNA, 0.05mM of dNTPs, 100 pmol of each primer and 1unit of Taq polymerase in 50µl of Taq buffer with MgCl2 (Boehringer Mannheim). The reaction was initiated at 94°C for 3.5 minutes followed by 30 cycles of 94°C for 1.5 minutes, 50°C for 1.5 minutes and 72°C for 3 minutes, with the final extension at 72°C for 10 minutes. The amplified 0.7kb fragment was digested with EcoRI and SpeI, and purified on 1.5% agarose gel. Obtained fragment was subjected to the ligation with pCZR134 linearized by EcoRI and SpeI. The resulting

plasmid was termed pCaXI1.

Gene synthesis of *Themomyces* xylanase gene

The *Themomyces* xylanase gene was synthesized to change the codon usage to have A or T as the preferred third letter of each codon. This is not a full optimization to the preferred codon usage of *Pichia methanolica*, as defined herein, but rather a first attempt to see if there is any advantage in optimizing the codons further. Two fragments XyC12 and XyC37 were synthesized by two-step PCR. The resulting designed xylanase encoding sequence is shown as SEQ ID 3.

Oligonucleotides used for the first PCR reaction were as follows:

For the synthesis of XyC12:

xyC1 (SEQ ID 4): atggttggttttactccagttgctttagctgctttagctgctactggtgctttagctttc ccagcaggtaatgctactgaattagaaaaaagacaaacaactccaaattctgaaggttgg.

xyC2r (SEQ ID 5): cataggtaccaccttctaaattagtataagtagcttgagcaccaccatca gaccaccaagaataataataaccatcatgccaaccttcagaatttggaga.

For the synthesis of XyC37:

xyC3 (SEQ ID 6): ggtacctatgaaatttcttggggtgatggtggtaatttagttggtggtaaaggttggaat ccaggtttaaatgcaagagctattcattttgaaggtgtttatcaaccaaatggtaattct.

xyC4r (SEQ ID 7): ctagtagttttacctaatctataaatagaaccatcacattcaacagtacctaaatcag tagcaccagaagatggatcataagtaccaaaattttcaacaatataatattcaactaatgga.

xyC5 (SEQ ID 8): ttctggtgctactgatttaggtactgttgaatgtgatggttctatttatagattaggtaa aactactagagttaatgcaccatctattgatggtactcaaactttcgatcaatattggtc.

xyC6r (SEQ ID 9): taataatgatcaccattaacattcaaaccagctctagcccaagcatcgaaatgac aaccagtttgaacagtaccagaagttcttttatcttgtctaacagaccaatattgatcgaaagtt.

xyC7 (SEQ ID 10): ttaatggtgatcattattatcaaattgttgcaactgaaggttatttctcttctggttatg ctagaattactgttgctgatgttggttaa.

Oligonucleotides used for the second PCR reaction were as follows:

For the synthesis of XyC12:

xyc12N (SEQ ID 11): ataagaatgcggccgcatggttggttttactccagt.

xyc12C (SEQ ID 12): tcataggtaccaccttctaaattagtataa.

For the synthesis of XyC37:

xyc37N (SEQ ID 13): ttagaaggtggtacctatgaaatttcttggggtgat.

xyc37C (SEQ ID 14): ttttccttttgcggccgcttaaccaacatcagcaac.

The first PCR reaction mixture contained 1pmol of each oligonucleotide, 0.35mM dNTP, 2.6 units of Expand™ polymerase in 50µl of provided buffer. Reaction started at 94°C for 1 minute followed by 6 cycles of 94°C for 0.5 minutes, 50°C for 5 minutes and 72°C for 1 minute with the final extension at 72°C for 10 minutes. Five µl of resulting first PCR reaction solution was used for the second PCR with 0.35mM dNTP, 100 pmol of each oligonucleotide and 2.6 units of Expand™ polymerase in 50µl of provided buffer. The reaction was carried out at 94°C for 3.5 minutes first followed by 30 cycles of 94°C for 1.5 minutes, 50°C for 1:5 minutes and 72°C for 3 minutes with the final extension at 72°C for 10 minutes. The amplified fragments were purified on 4% agarose gel then ligated into pT7blue T-vector resulting pTXyC12-7 and pTXyC37-715 respectively. They were confirmed to have no PCR error by sequencing. A KpnI fragment of around 200bp was isolated from pTXyC12-7 then ligated with a 3.7kb of KpnI fragment from pTXyC37-715. The resulting plasmid pTXyC17-1 comprises the sequence of the codon-optimized synthesized *Themomyces* xylanase gene (SEQ ID 3).

Construction of expression plasmid pCaXS3

The synthesized *Thermomyces* xylanase gene was amplified by PCR using the following primers;

a) (SEQ ID 15): CCGGAATTCATGGTTGGTTTTACTCCAGTTGCT.

b) (SEQ ID 16): GGACTAGTTTAACCAACATCAGCAACAGTAAT.

The reaction mixture comprised 10 ng of synthesized xylanase gene, 0.35 mM of dNTP, 100 pmol of each primer and 2.6 units of Expand™ polymerase in 50µl of provided buffer with MgCl2. The PCR reaction was initiated at 94°C for 3.5 minutes followed by 30 cycles of 94°C for 1.5 minutes, 55°C for 1.5 minutes and 68°C for 3 minutes with the final extension at 72°C for 10 minutes. The amplified gene was digested with EcoRI and SpeI, and ligated into pCZR134 cut by EcoRI and SpeI beforehand and the resulting plasmid was termed pCaXS3.

Transformation of *P. methanolica* with pCaXI1 and pCaXS3

The expression plasmids pCaXI1 and pCaXS3 were digested with NotI and 6.2kb fragments were used for the transformation of *P. methanolica* PMAD16 to achieve a Ade+ phenotype. Transformed cells were grown on SC glucose (2% glucose, 0.01% tryptophan, 0.02% threonine, 0.5% casamino acid, in basal salt without amino acids (per litre, 6.68g yeast nitrogen base w/o amino acid (Difco), 10g succinic acid, 6g NaOH, and 2% Agar Noble (Difco) containing 1.2Msorbitol). Rapidly growing white colonies were isolated as stable transformants.

Obtained transformants were grown on YPM agar (20g/l of Pepton, Difco, 10g/l of yeast extract, Difco, 20g/l of Agar Noble, Difco and 20ml/l of methanol) containing

0.1% AZCL™-xylan (Megazyme™) for 3-5 days at 30°C. Positive transformants were identified as the colonies surrounded by a blue halo.

Xylanase producing cells were inoculated in YPD medium (20g/l of Pepton, Difco, 10g/l of yeast extract, Difco and 20g/l of glucose) in a 500 ml shaking flask, and cultivated under vigorous aeration overnight at 30°C. One ml of methanol was added to each shaking flask and it was cultivated one more day. Cells were removed by centrifugation and xylanase activity in the supernatant was measured using 0.5% Azo-Wheat Arabinoxylan (Megazyme™) as substrate. The measurement reaction was employed at 30°C, pH 6 for 30 minutes and released colour was measured as the absorbance at 585nm. The amount of enzyme protein was estimated from the activity. The results are shown below; our conclusion to this first attempt is that the codon optimization did indeed have a significant positive effect on the obtainable yield from a *Pichia* cell. However, since we did not fully optimize the codons of the xylanase encoding sequence to match the preferred codon usage of *P. methanolica*, it may be possible to obtain even greater yields by adjusting the sequence thoroughly.

Strain	Plasmid	Gene	Yield in shaking flask
PXI242	pCaXI1	intact gene	12 mg/l
PXS5	pCaXS3	synthesized gene	20 mg/l

A transformant carrying the synthesized *Thermomyces* xylanase gene was inoculated in two flasks with 110ml of YEPD (10g/l Bacto yeast extract, 20g/l Bacto peptone and 20g/l glucose) in 500ml shaking flasks then cultivated under vigorous aeration condition at 30°C for 16-18 hours. Most of the cells were used for the subsequent inoculation of the fermentor. The fermentation vessel was filled with 2 liters of media containing 43g $(NH_4)_2SO_4$, 5.2g K_2HPO_4, 19g KH_2PO_4, 0.8g $FeSO_4.7H_2O$, 2.3g sodium citrate with 2aq. After autoclaving and cooling down to 30°C, 94ml of 60%

glucose, 40ml of 1M MgSO4.7H2O, 4ml of 1M CaCl2, 40ml of trace metal solution (containing, per litter, 27.8g Fe SO4.7H2O, 0.5g CuSO4.5H2O, 1.09g ZnCl2, 1.35g MnSO4.H2O, 0.48g CoCl2.6H2O, 0.24g Na2MoO4.2 H2O, 0.5g H3BO3, 0.08g KI, 0.26g NiSO4.6 H2O, 0.56g thiamine hydrochloride, 5mg biotin, 1ml H2SO4) and 4ml of vitamin solution (containing, per litter, 47g inositol, 25g calcium pantothenate, 1.3g pyridoxine hydrochloride, 5.6g thiamine hydrochloride, 0.1g biotin) were added. The pH of the fermentation media was adjusted to 5.0 and during the fermentation pH was controlled automatically with 25% ammonia water and 10% H_3PO_4. Aeration was provided as compressed air at a flow rate of 3 litters per minute initially and agitation at the rate of 100 rpm. After inoculation the agitation was controlled automatically from 100 rpm to 1000 rpm to keep the dissolved oxygen level at minimum 30% of saturation. Following exhaustion of glucose after inoculation, a glucose feed was initiated at 12 hours at the rate of 0.27% glucose/hour. The feed was increased stepwise to 0.38% glucose/hour at 18 hours and to 0.5% glucose/hour at 30 hours. After 40 hours, methanol induction of the AUG1 promoter was started by methanol feed at the rate of 0.3% methanol/hour, in parallel the glucose feed rate was reduced to 0.38% glucose/hour. Finally mixed feeding (0.6% methanol/hour, 0.27% glucose/hour) was run for the rest of fermentation. Increasing methanol feeding elevated the enzyme productivity. The yield of *Thermomyces* xylanase was calculated to be 750 mg/l with the 110g/kg of dried cells at 170 hours.

Example 2

cDNA clone of *A.fumigatus* phytase gene

The sequence of the phytase gene from *Aspergillus fumigatus* ATCC 13073 was published previously (EP 0897010). This phytase contains an amino acid mutation Q27L, and the phytase of the invention comprises a further mutation at the second amino acid,

V2I, due to the introduction of a restriction enzyme cleavage site around the translation initiation codon.

Construction of expression plasmid pCaPhy3

A cDNA of *A. fumigatus* phytase gene was amplified by PCR using the primers shown below;

A1) (SEQ ID 17): CCGGAATTCATGATTACTCTGACTTTCCTGCTTTCGGCGG CGTATCTGCTTTCTGGTAGAGTGTCTGCGGCACCTAGTTCTGCTGGCTC.

B1) (SEQ ID 18): CTAGACTAGTCTAGTCAACTAAAGCACTCTCCCCA.

The reaction mixture comprising 10 ng of cDNA clone of *A.fumigatus* phytase, 0.35 mM dNTPs, 100 pmol of each primers and 2.6 units of Expand™ polymerase in 50µl of provided buffer with MgCl2, was incubated at 94°C for 2 minutes followed by 10 cycles of 94°C for 10 seconds, 55°C for 30 seconds, 68°C for 45 seconds. After that another 20 cycles were performed as 94°C for 10 seconds, 55°C for 30 seconds, 68°C for 45 plus n x 20 seconds (n: cycle number).

The final extension was at 68°C for 7 minutes. Amplified 1.3kb fragment was purified on 1.2% agarose gel and digestion with EcoRI and SpeI. It was used for the ligation with pCZR134 digested with the same restriction enzymes and the obtained plasmid was termed pCaPhy3.

Construction of the expression plasmid pCPS40

DNA encoding the *A. fumigatus* phytase was synthesized, the sequence was designed based on the codon usage of two genes from *Pichia methanolica*, i.e. ADE1 and AUG1. Codon usage of designed sequence is shown in table 1.

This designed codon optimized phytase DNA sequence is shown in (SEQ ID 19).

Table 1

Amino acid	Codon	P.methanolica ADE1	P.methanolica AUG1	Original phytase gene	Optimized phytase gene
F	TTT	9	10	12	12
	TTC	6	19	15	15
L	TTA	17	16	0	31
	TTG	14	19	10	10
	CTT	1	2	7	0
	CTC	2	0	5	0
	CTA	7	3	4	4
	CTG	2	0	19	0
I	ATT	19	22	8	9
	ATC	18	12	6	6
	ATA	5	0	1	0
M	ATG	17	17	6	6
V	GTT	29	36	6	25
	GTC	12	4	4	4
	GTA	0	0	3	0
	GTG	3	0	16	0
S	TCT	10	20	6	15
	TCC	5	10	10	17
	TCA	9	11	3	16
	TCG	4	0	13	0
	AGT	2	0	9	0
	AGC	3	0	7	0
P	CCT	7	1	7	7
	CCC	1	0	4	0
	CCA	19	44	4	15
	CCG	5	0	7	0
T	ACT	22	31	11	21
	ACC	6	11	9	9
	ACA	4	2	1	1
	ACG	2	0	10	0
A	GCT	27	40	9	25
	GCC	14	7	12	18
	GCA	7	1	6	0
	GCG	3	0	16	0
Y	TAT	5	9	4	4
	TAC	9	20	14	14
H	CAT	8	4	4	4
	CAC	7	18	4	4
Q	CAA	18	12	2	15
	CAG	2	0	13	0
N	AAT	11	2	6	0

	AAC	13	30	13	19
K	AAA	15	7	1	1
	AAG	17	31	25	25
D	GAT	15	30	8	8
	GAC	13	18	17	17
E	GAA	20	33	4	20
	GAG	7	3	16	0
C	TGT	9	11	2	10
	TGC	1	2	8	0
W	TGG	2	10	5	5
R	CGT	2	1	1	1
	CGC	0	0	7	0
	CGA	0	0	3	0
	CGG	0	0	6	0
	AGA	14	31	2	18
	AGG	2	0	1	1
G	GGT	26	53	3	33
	GGC	11	1	16	0
	GGA	3	0	6	0
	GGG	2	0	8	0
Stop	TAA	0	0	0	0
	TAG	0	1	0	0
	TGA	1	0	1	1
Total		544	665	466	466

Seventy oligonucleotides (SEQ ID's 20-89) were prepared for the construction of the codon optimized phytase gene, and subjected to PCR using Expand ™ Long Template (Boehringer Mannheim) for assembly. The reaction mixture comprised 125 pmol of each of the seventy oligonucleotides shown in SEQ ID's 20-89, 0.2 mM dNTPs and 2.6 units of Expand ™ polymerase in 50µl of provided buffer. The reaction was employed as 55 cycles of 94°C for 30 seconds, 52°C for 30 seconds and 68°C for 30 seconds. The second PCR was carried out using the resulting first reaction mixture as template and the following set of primers:

c) (SEQ ID 90): CGGAATTCATGATTACTTTAACTTTCTTATTATCAGCTTATTTA.

d) (SEQ ID 91): GGACTAGTTCAAGAAAAACATTCACCCC.

The reaction mixtures comprised 0.5µl, 1µl or 2µl of the first PCR mixture, 30 pmol of each primer, 0.2 mM dNTPs and 2.6 units of Expand ™ polymerase in 50µl of provided buffer. The reaction was employed as 25 cycles of 94°C for 30 seconds, 50°C for 30 seconds and 68°C for 60 + 10 x n seconds (n: cycle number). An amplified 1.3kb fragment was purified with QIAquick and cloned into pT7blue T-vector. Two independent plasmids pTPhy5 and pTPhy8 were selected for further construction, they had a couple of PCR generated sequence errors within the HindIII–SpeI region of pTPhy5 and the EcoRI-HindIII region of pTPhy8.

Each plasmid was digested with HindIII, and a 3.4kg fragment from pTPhy5 and a 0.7kb fragment from pTPhy8 were isolated and ligated together, resulting in the plasmid pTPhy58, which comprises the entire sequence of the synthetic *A. fumigatus* phytase gene.

The plasmid pTPhy58 was digested with EcoRI and SpeI, and a 1.3kb fragment was isolated and ligated into the expression vector pCZR134. The resulting phytase expression plasmid was designated pCPS40.

The entire sequence of the synthetic codon optimized *A. fumigatus* phytase gene (SEQ ID No. 19) is comprised in the strain *E. coli* NN049526, which was deposited with DSMZ 01-March-2000 with the accession number DSM 13352. The deposit was made by Novo Nordisk A/S and was later assigned to Novozymes A/S.

Transformation of *P. methanolica* with pCaPhy3 and pCPS40

The expression plasmids pCaPhy3 and pCPS40 were digested with NotI and 6.7kb fragments were used for the transformation of *P. methanolica* PMAD16. Transformed cells were grown on minimal medium SC glucose containing 1.2 M sorbitol. White colonies growing rapidly were isolated as stable transformants. The obtained transformants were subsequently grown on YPM agar containing 0.1% phytic acid

(Sigma P8810) at 30°C overnight, then 0.1M CaCl2 was poured onto the grown agar plate to detect phytase production. Positive transformants were detected as colonies surrounded by a clear halo.

Phytase producing strains were inoculated in 50ml of YPD medium in 500ml of shaking flasks and cultivated at 30°C with vigorous aeration. After overnight cultivation, 1ml of methanol was added to each shaking flask and the cultivation was continued one more day. Cells were removed by centrifugation and phytase activity in the supernatant was measured using 0.15% sodium phytate in 100mM sodium acetate buffer pH 5.5 as the substrate. The reaction was carried out at room temperature for 15 minutes, and liberated phosphate ions were quantified by colorization with 1% ammonum molybdate, 0.26M FeSO4 in 3.2% H2SO4 and absorbance was measured at 650nm. The amount of enzyme protein was calculated from the activity based on the known specific activity per mg enzyme protein. The results are shown in Table 2.

Table 2.

Strain	Plasmid	Used gene	Yield in shaking flask
PPi 51	pCaPhy3	cDNA	1 mg/l
PPs237	pCPS40	codon optimized	20 mg/l

The transformants mentioned above were cultivated as described in example 1. The yield of *A. fumigatus* phytase from a transformant generated with pCaPhy3, carrying non codon optimized cDNA of the phytase gene, was calculated to be 310 mg/l with 120 g/kg of dried cells at 170 hours of fermentation, whereas the yield from the transformant with a codon-optimised gene was estimated to be 3.5 g/l with 115 g/kg of dried cells at 190 hours of fermentation.

Thus the yield from the codon optimized transformant was approximately 10 times higher than what could be obtained from the transformant carrying the original

cDNA phytase gene.

Example 3

Expression of the *A.fumigatus* phytase from the GAP-promoter

A synthesized *A.fumigatus* phytase gene was obtained from pCPS40 as described in Example 2 by digesting with the restriction enzymes EcoRI and SpeI. The isolated 1.3kb fragment was ligated into the pGAP vector, which was initially pre-digested with the same enzymes, the resulting plasmid was termed pGAPphy. The plasmid pGAPphy was then digested with NotI and the 6.9kb fragment was used for the transformation of *P.methanolica* PMAD16. Transformants were isolated as described in Example 2.

Obtained transformants were cultivated in 5ml of YPD medium at 30°C overnight with vigorous aeration, then 1ml of grown culture was transferred to fresh 100ml of YPD medium in a shaking flask. After one-day cultivation at 30°C, 2ml of 20% glucose solution was added to each flask and cultivation was continued for one more day. Enzyme activity of the culture broth, from which the grown cells were removed, was measured to estimate the enzyme productivity. The estimated phytase productivity was 47mg/l.

A transformant carrying a synthesized *Aspergillus fumigatus* phytase gene was fermented in 5L jar-fermenter. The fermentation vessel was filled with 2 liters of medium containing 43g (NH4)2SO4, 5.2g K2HPO4, 19g KH2PO4, 0.8g Fe SO4.7H2O, 2.3g sodium citrate with 2aq. After autoclaving and cooling down to 30°C, 94ml of 60% glucose, 40ml of 1M MgSO4.7H2O, 4ml of 1M CaCl2, 40ml of trace metal solution and 4ml of vitamin solution were added. The pH of fermentation media was adjusted to 5.0 and during the fermentation it was controlled automatically above 5.0 with 25% ammonia water. Aeration was provided as compressed air at a flow rate of 3 litres per minute initially and agitation at the rate of 100 rpm. After inoculation the agitation was controlled automatically from 100 rpm to 1000 rpm to keep the dissolved oxygen level at minimum 30% of saturation. Around 11 hours after inoculation, 1100ml of 60% glucose was supplemented. The yield of *Aspergillus fumigatus* phytase at 50 hours was calculated to 1000 mg/l with the 65 g/kg of dried cells.

CLAIMS

1. A *Pichia* cell comprising at least one copy of a heterologous nucleotide sequence encoding a polypeptide of interest, wherein the codon usage of said sequence has been adjusted to match the preferred codon usage of *P. methanolica*, as defined herein.

2. The cell of claim 1, which is a *Pichia methanolica* cell.

3. The cell of claim 1 or 2, wherein the nucleotide sequence encodes an enzyme.

4. The cell of claim 1 - 3, wherein the nucleotide sequence encodes an enzyme having phytase or xylanase activity.

5. The cell of claim 1, wherein the nucleotide sequence comprises a DNA sequence at least 90% identical to the sequences shown in SEQ ID No. 3, SEQ ID No. 19, or to the phytase encoding sequence comprised in a cell of DSM 13352.

6. The cell of claims 1 – 5, wherein the nucleotide sequence is present on an extrachromosomal DNA construct, preferably a plasmid.

7. The cell of claims 1 – 5, wherein the nucleotide sequence is integrated in the cell's chromosome.

8. The cell of claims 1 - 7, wherein the nucleotide sequence is transcribed from a promoter of a methanol-inducible *P. methanolica* gene.

9. The cell of claim 8, wherein the promoter is a native promoter of a *P. methanolica*

gene encoding an enzyme selected from the group consisting of alcohol oxidase, dihydroxyacetone synthase, formate dehydrogenase, and catalase.

10. The cell of claims 1 - 9, wherein the polypeptide comprises its native original secretion signal sequence.

11. An isolated DNA construct comprising at least one copy of a nucleotide sequence encoding a polypeptide heterologous to *Pichia methanolica*, wherein the codon usage of said sequence has been adjusted to match the preferred codon usage of *P. methanolica*, as defined herein.

12. The DNA construct of claim 11, wherein the nucleotide sequence encodes an enzyme.

13. The DNA construct of claim 11, wherein the nucleotide sequence encodes an enzyme having phytase or xylanase activity.

14. The DNA construct of claim 11, wherein the nucleotide sequence comprises a DNA sequence at least 90% identical to the sequences shown in SEQ ID No. 3, SEQ ID No. 19, or to the phytase encoding sequence comprised in a cell of DSM 13352.

15. The DNA construct of claims 11 – 14, wherein the nucleotide sequence is transcribed from a promoter of a methanol-inducible *P. methanolica* gene.

16. The DNA construct of claim 15, wherein the promoter is a native promoter of a *P. methanolica* gene encoding an enzyme selected from the group consisting of alcohol

oxidase, dihydroxyacetone synthase, formate dehydrogenase, and catalase.

17. The DNA construct of claims 11 – 16, wherein the nucleotide sequence encodes a polypeptide comprising its native original secretion signal sequence.

18. A method of producing a polypeptide of interest in a *Pichia* cell, where the polypeptide is encoded by a nucleotide sequence heterologous to *P. methanolica*, said method comprising the steps of:

a) adjusting the codon usage of the sequence to match the preferred codon usage of *P. methanolica*, as defined herein; and

b) cultivating a *Pichia* cell comprising at least one copy of the codon usage adjusted sequence of step a) under appropriate growth conditions to express the sequence and achieve production of the polypeptide of interest.

19. The method of claim 18, wherein the *Pichia* cell is a *Pichia methanolica* cell.

20. The method of claim 18 or 19, wherein the nucleotide sequence encodes an enzyme.

21. The method of claim 18 or 19, wherein the nucleotide sequence encodes an enzyme having phytase or xylanase activity.

22. The method of claim 18 or 19, wherein the nucleotide sequence comprises a DNA sequence at least 90% identical to the sequences shown in SEQ ID No. 3, SEQ ID No. 19, or to the phytase encoding sequence comprised in a cell of DSM 13352.

23. The method of claims 18 - 22, wherein the polypeptide is produced in a yield at least three times higher than what is obtained from an otherwise identical cell comprising a sequence not adjusted to the preferred codon usage of *P. methanolica*, as defined herein. (Preferably at least five times higher, preferably seven, more preferably eight, even more preferably nine, and most preferably ten times higher.)

24. The method of claims 18 – 23, wherein the nucleotide sequence is transcribed from a promoter of a methanol-inducible *P. methanolica* gene.

25. The method of claim 24, wherein the promoter is a native promoter of a *P. methanolica* gene encoding an enzyme selected from the group consisting of alcohol oxidase, dihydroxyacetone synthase, formate dehydrogenase, and catalase.

26. The method of claim 24 or 25, wherein methanol is added to the cell during the cultivating step, whereby the methanol-inducible promoter is induced.

27. The method of claims 18 - 26, wherein the nucleotide sequence encodes a polypeptide comprising its native original secretion signal sequence.

SEQUENCE LISTING

<110> Novozymes A/S

<120> High yield production of polypeptides in Pichia methanolica

<130> 6125.200-US

<140>
<141>

<160> 91

<170> PatentIn Ver. 2.1

<210> 1
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyN

<400> 1
ccggaattca tggtcggctt tacccccgtt 30

<210> 2
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer XyC

<400> 2
aaggaaaaaa gcggccgcac tagtttagcc cacgtcagca ac 42

<210> 3
<211> 678
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Codon optimized xylanase encoding sequence.

<400> 3
atggttggtt ttactccagt tgctttagct gctttagctg ctactggtgc tttagctttc 60
ccagcaggta atgctactga attagaaaaa agacaaacaa ctccaaattc tgaaggttgg 120
catgatggtt attattattc ttggtggtct gatggtggtg ctcaagctac ttatactaat 180
ttagaaggtg gtacctatga aatttcttgg ggtgatggtg gtaatttagt tggtggtaaa 240
ggttggaatc caggtttaaa tgcaagagct attcattttg aaggtgttta tcaaccaaat 300
ggtaattctt atttagctgt ttatggttgg actagaaatc cattagttga atattatatt 360
gttgaaaatt ttggtactta tgatccatct tctggtgcta ctgatttagg tactgttgaa 420
tgtgatggtt ctatttatag attaggtaaa actactagag ttaatgcacc atctattgat 480
ggtactcaaa ctttcgatca atattggtct gttagacaag ataaaagaac ttctggtact 540

gttcaaactg gttgtcattt cgatgcttgg gctagagctg gtttgaatgt taatggtgat 600
cattattatc aaattgttgc aactgaaggt tatttctctt ctggttatgc tagaattact 660
gttgctgatg ttggttaa 678

<210> 4
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer XyC1

<400> 4
atggttggtt ttactccagt tgctttagct gctttagctg ctactggtgc tttagctttc 60
ccagcaggta atgctactga attagaaaaa agacaaacaa ctccaaattc tgaaggttgg 120

<210> 5
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer XyC2r

<400> 5
cataggtacc accttctaaa ttagtataag tagcttgagc accaccatca gaccaccaag 60
aataataata accatcatgc caaccttcag aatttggaga 100

<210> 6
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer XyC3

<400> 6
ggtacctatg aaatttcttg gggtgatggt ggtaatttag ttggtggtaa aggttggaat 60
ccaggtttaa atgcaagagc tattcatttt gaaggtgttt atcaaccaaa tggtaattct 120

<210> 7
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer XyC4r

<400> 7
ctagtagttt tacctaatct ataaatagaa ccatcacatt caacagtacc taaatcagta 60
gcaccagaag atggatcata agtaccaaaa ttttcaacaa tataatattc aactaatgga 120

<210> 8
<211> 120

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyC5

<400> 8
ttctggtgct actgatttag gtactgttga atgtgatggt tctatttata gattaggtaa 60
aactactaga gttaatgcac catctattga tggtactcaa actttcgatc aatattggtc 120

<210> 9
<211> 120
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyC6r

<400> 9
taataatgat caccattaac attcaaacca gctctagccc aagcatcgaa atgacaacca 60
gtttgaacag taccagaagt tcttttatct tgtctaacag accaatattg atcgaaagtt 120

<210> 10
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyC7

<400> 10
ttaatggtga tcattattat caaattgttg caactgaagg ttatttctct tctggttatg 60
ctagaattac tgttgctgat gttggttaa 89

<210> 11
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyc12N

<400> 11
ataagaatgc ggccgcatgg ttggttttac tccagt 36

<210> 12
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyc12C

<400> 12
tcataggtac caccttctaa attagtataa 30

<210> 13
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyc37N

<400> 13
ttagaaccat cta tttca aatttcttgg ggtgat 36

<210> 14
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer xyc37C

<400> 14
ttttccttta tttttttt aatcaacatc agcaac 36

<210> 15
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer a)

<400> 15
ccggaattca tgattggttt tactccagtt gct 33

<210> 16
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer b)

<400> 16
ggactagttt aaccaacatc agcaacagta at 32

<210> 17
<211> 89
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer A1

<400> 17

ccggaattca tgattacttt gacttttcctg ctttcggcgg cgtatctgct ttctggtaga 60
gtgtctgcgg cacctagttc tgctggctc 89

<210> 18
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer B1.

<400> 18
ctagactagt ctagtcaact aaagcactct cccca 35

<210> 19
<211> 1398
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Pichia methanolica codon optimized phytase encoding sequence.

<400> 19
atgattactt taactttctt attatcagct gcttatttat tatctggtag agtttctgct 60
gccccttctt ctgctggttc caagtcctgt gatactgttg acttaggtta ccaatgttcc 120
cctgctactt ctcatctatg gggtttatac tcaccattct tttcattaga agacgaatta 180
tccgtttcat ctaagttacc aaaggattgt agaatcacct tggttcaagt tctatcaaga 240
catggtgcta gatacccaac ctcctccaag tccaaaaagt ataagaagtt agttactgct 300
atccaagcca acgccaccga cttcaagggt aagtttgcct ttttgaagac ttacaactat 360
actttaggtg ctgatgactt aactccattt ggtgaacaac aattagttaa ctcaggtatc 420
aagttctacc aaaggtacaa ggctttagct agatctgttg ttccatttat tagagcctca 480
ggttcagaca gagttattgc ttcaggtgaa aagttcatcg aaggtttcca acaagctaag 540
ttagctgatc ctggtgctac taacagagcc gctccagcta tttctgttat tattccagaa 600
tccgaaactt tcaacaacac tttagaccac ggtgtttgta ctaagtttga agcttctcaa 660
ttaggtgatg aagttgctgc caacttcact gctttatttg ccccagacat cagagctaga 720
gccgaaaagc atttacctgg tgttacttta acagacgaag acgtcgtttc tctaatggac 780
atgtgttcat ttgatactgt tgctagaacc tccgacgcct ctcaattatc accattctgt 840
caattattca ctcacaacga atggaagaag tacaactact tacaatcctt gggtaagtac 900
tacggttacg gtgccggtaa ccctttaggt ccagctcaag gtattggttt caccaacgaa 960
ttaattgcca gattgactcg ttcaccagtt caagaccaca cctccactaa ctcaactcta 1020
gtctccaacc cagccacctt cccattgaac gctaccatgt acgtcgactt ttcacacgac 1080
aactccatgg tttccatctt ctttgccttg ggtttataca acggtactga accattgtcc 1140
agaacctcag ttgaatccgc caaggaattg gatggttatt ctgcctcctg ggttgttcct 1200
ttcggtgcta gagcctactt cgaaactatg caatgtaagt cagaaaagga acctttagtt 1260
agagctttga ttaacgacag agttgttcca ttacatggtt gtgatgttga caagttaggt 1320
agatgtaagt taaacgactt tgtcaagggt ttgtcttggg ccagatctgg tggtaactgg 1380
ggtgaatgtt tttcttga 1398

<210> 20
<211> 40
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer phy1.

<400> 20
atgattactt taacttttctt attatcagct gcttatttat 40

<210> 21
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy2r

<400> 21
agcagaaact ctaccagata ataaataagc agctgataat 40

<210> 22
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy3

<400> 22
tatctggtag agtttctgct gccccttctt ctgctggttc 40

<210> 23
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy4r

<400> 23
caacagtatc acaggacttg gaaccagcag aagaaggggc 40

<210> 24
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy5

<400> 24
caagtcctgt gatactgttg acttaggtta ccaatgttcc 40

<210> 25
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy6r

<400> 25
catagatgat [illegible] ggaacattgg taacctaagt 40

<210> 26
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy7

<400> 26
cctgctactt [illegible] gattttatac tcaccattct 40

<210> 27
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy8r

<400> 27
taattcgtct tctaatgaaa agaatggtga gtataaaccc 40

<210> 28
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy9

<400> 28
tttcattaga agacgaatta tccgtttcat ctaagttacc 40

<210> 29
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy10r

<400> 29
aggtgattct acaatccttt ggtaacttag atgaaacgga 40

<210> 30
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy11

<400> 30
aaaggattgt agaatcacct tggttcaagt tctatcaaga 40

<210> 31
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy12r

<400> 31
gttgggtatc tagcaccatg tcttgataga acttgaacca 40

<210> 32
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy13

<400> 32
catggtgcta gatacccaac ctcctccaag tccaaaaagt 40

<210> 33
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy14r

<400> 33
agcagtaact aacttcttat acttttgga cttggaggag 40

<210> 34
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy15

<400> 34
ataagaagtt agttactgct atccaagcca acgccaccga 40

<210> 35
<211> 40
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy16r

<400> 35
aggcaaactt accettgaag tcggtggcgt tggcttggat 40

<210> 36
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy17

<400> 36
cttcaagggt aagtttgcct ttttgaagac ttacaactat 40

<210> 37
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy18r

<400> 37
aagtcatcag cacctaaagt atagttgtaa gtcttcaaaa 40

<210> 38
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy19

<400> 38
actttaggtg ctgatgactt aactccattt ggtgaacaac 40

<210> 39
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy20r

<400> 39
gatacctgag ttaactaatt gttgttcacc aaatggagtt 40

<210> 40
<211> 40

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy21

<400> 40
aattagttaa ctcaggtatc aagttctacc aaaggtacaa 40

<210> 41
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy23

<400> 41
caacagatct agctaaagcc ttgtaccttt ggtagaactt 40

<210> 42
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy23

<400> 42
ggctttagct agatctgttg ttccatttat tagagcctca 40

<210> 43
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy24r

<400> 43
gcaataactc tgtctgaacc tgaggctcta ataaatggaa 40

<210> 44
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy25

<400> 44
ggttcagaca gagttattgc ttcaggtgaa aagttcatcg 40

<210> 45

<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy26r

<400> 45
cttagcttgt tggaaacctt cgatgaactt ttcacctgaa 40

<210> 46
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy27

<400> 46
aaggtttcca acaagctaag ttagctgatc ctggtgctac 40

<210> 47
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy28r

<400> 47
tagctggagc ggctctgtta gtagcaccag gatcagctaa 40

<210> 48
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy29

<400> 48
taacagagcc gctccagcta tttctgttat tattccagaa 40

<210> 49
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy30r

<400> 49
gtgttgttga aagtttcgga ttctggaata ataacagaaa 40

<210> 50
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy31

<400> 50
tccgaaactt tcaacaacac tttagaccac ggtgtttgta 40

<210> 51
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy32r

<400> 51
ttgagaagct tcaaacttag tacaaacacc gtggtctaaa 40

<210> 52
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy33

<400> 52
ctaagtttga agcttctcaa ttaggtgatg aagttgctgc 40

<210> 53
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy34r

<400> 53
caaataaagc agtgaagttg gcagcaactt catcacctaa 40

<210> 54
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy35

<400> 54
caacttcact gctttatttg ccccagacat cagagctaga 40

<210> 55
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy36r

<400> 55
ccaggtaaat gcttttcggc tctagctctg atgtctgggg 40

<210> 56
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy37

<400> 56
gccgaaaagc atttacctgg tgttacttta acagacgaag 40

<210> 57
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy38r

<400> 57
gtccattaga gaaacgacgt cttcgtctgt taaagtaaca 40

<210> 58
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy39

<400> 58
acgtcgtttc tctaatggac atgtgttcat ttgatactgt 40

<210> 59
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy40r

<400> 59
aggcgtcgga ggttctagca acagtatcaa atgaacacat 40

<210> 60
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy41

<400> 60
tgctagaacc tccgacgcct ctcaattatc accattctgt 40

<210> 61
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy42r

<400> 61
tcgttgtgag tgaataattg acagaatggt gataattgag 40

<210> 62
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy43

<400> 62
caattattca ctcacaacga atggaagaag tacaactact 40

<210> 63
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy44r

<400> 63
gtacttaccc aaggattgta agtagttgta cttcttccat 40

<210> 64
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy45

<400> 64

tacaatcctt gggtaagtac tacggttacg gtgccggtaa 40

<210> 65
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy46r

<400> 65
cttgagctgg acctaaaggg ttaccggcac cgtaaccgta 40

<210> 66
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy47

<400> 66
ccctttaggt ccagctcaag gtattggttt caccaacgaa 40

<210> 67
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy48r

<400> 67
cgagtcaatc tggcaattaa ttcgttggtg aaaccaatac 40

<210> 68
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy49

<400> 68
ttaattgcca gattgactcg ttcaccagtt caagaccaca 40

<210> 69
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy50r

<400> 69
tagagttgaa tttaaggaag tgtggtcttg aactggtgaa 40

<210> 70
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy51

<400> 70
cctccactta tttaaaaaca gtctccaacc cagccacctt 40

<210> 71
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy52r

<400> 71
acatggtagc ttt attaag aaggtggctg ggttggagac 40

<210> 72
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy53

<400> 72
cccattgaac gctaccatgt acgtcgactt ttcacacgac 40

<210> 73
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy54r

<400> 73
aagatggaaa ccatggagtt gtcgtgtgaa aagtcgacgt 40

<210> 74
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy55

<400> 74
aactccatgg tttccatctt ctttgccttg ggtttataca 40

<210> 75
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy56r

<400> 75
ggacaatggt tcagtaccgt tgtataaacc caaggcaaag 40

<210> 76
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy57

<400> 76
acggtactga accattgtcc agaacctcag ttgaatccgc 40

<210> 77
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy58r

<400> 77
aataaccatc caattccttg gcggattcaa ctgaggttct 40

<210> 78
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy59

<400> 78
caaggaattg gatggttatt ctgcctcctg ggttgttcct 40

<210> 79
<211> 40
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer phy60r

<400> 79
aagtaggctc tagcaccgaa aggaacaacc caggaggcag 40

<210> 80
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy61

<400> 80
ttcggtgcta gagcctactt cgaaactatg caatgtaagt 40

<210> 81
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy62r

<400> 81
aactaaaggt tccttttctg acttacattg catagtttcg 40

<210> 82
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy63

<400> 82
cagaaaagga acctttagtt agagctttga ttaacgacag 40

<210> 83
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy64r

<400> 83
aaccatgtaa tggaacaact ctgtcgttaa tcaaagctct 40

<210> 84
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy65

<400> 84
agttgttttt ttatatttt gtgatgttga caagttaggt 40

<210> 85
<211> 4
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy66r

<400> 85
aagtcgtttt attttattt atttaacttg tcaacatcac 40

<210> 86
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy67

<400> 86
agatgtaagt ttaacgattt tgtcaagggt ttgtcttggg 40

<210> 87
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy68r

<400> 87
ccagttacca ccagatctgg cccaagacaa acccttgaca 40

<210> 88
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy69

<400> 88
ccagatctgg tggtaactgg ggtgaatgtt tttcttga 38

<210> 89
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer phy70r

<400> 89
tcaagaaaaa cattcacc 18

<210> 90
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer c)

<400> 90
cggaattcat gattacttta actttcttat tatcagctta ttta 44

<210> 91
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer d)

<400> 91
ggactagttc aagaaaaaca ttcacccc 28

6125.200-US

1

PCT 6125.204-WO

Original (for **SUBMISSION**) - printed on 09.03.2001 09:39:29 AM

0-1	Form - PCT/RO/134 (EASY) Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)	
0-1-1	Prepared using	PCT-EASY Version 2.91 (updated 01.01.2001)
0-2	International Application No.	
0-3	Applicant's or agent's file reference	6125.204-WO

1	The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:	
1-1	page	4
1-2	line	10
1-3	Identification of Deposit	
1-3-1	Name of depositary institution	DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
1-3-2	Address of depositary institution	Mascheroder Weg 1b, D-38124 Braunschweig, Germany
1-3-3	Date of deposit	01 March 2000 (01.03.2000)
1-3-4	Accession Number	DSMZ 13352
1-4	Additional Indications	NONE
1-5	Designated States for Which Indications are Made	all designated States
1-6	Separate Furnishing of Indications These indications will be submitted to the International Bureau later	NONE

FOR RECEIVING OFFICE USE ONLY

0-4	This form was received with the international application: (yes or no)	Yes
0-4-1	Authorized officer	

FOR INTERNATIONAL BUREAU USE ONLY

0-5	This form was received by the international Bureau on:	
0-5-1	Authorized officer	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 01/00154

A. CLASSIFICATION OF SUBJECT MATTER

IPC7: C12N 1/19, C12N 15/81
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC7: C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SE,DK,FI,NO classes as above

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0911404 A2 (PHONE PULENC AGRO), 28 April 1999 (28.04.99) --	1-27
X	WO 9966040 A1 (ZYMOGENETICS, INC.), 23 December 1999 (23.12.99), page 22, column 16 - page 23, line 12; page 46, column 12 - line 15 --	1-27
A	WO 9717450 A2 (ZYMOGENETICS, INC.), 15 May 1997 (15.05.97) --	1-27
A	WO 9632472 A1 (NOVO NORDISK A/S), 17 October 1996 (17.10.96), claim 7 --	1-27

[X] Further documents are listed in the continuation of Box C. [X] See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 July 2001	30-07-2001
Name and mailing address of the ISA Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. + 46 8 666 02 86	Authorized officer Patrick Andersson/BS Telephone No. + 46 8 782 25 00

INTERNATIONAL SEARCH REPORT

International application No.
PCT/DK 01/00154

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 9949022 A1 (NOVO NORDISK A/S), 30 Sept 1999 (30.09.99), figure 8	1-27

--

INTERNATIONAL SEARCH REPORT

Information on patent family members

02/07/01

International application No.
PCT/DK 01/00154

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0911404 A2	28/04/99	FR 2768748 A,B	26/03/99
		US 6180363 B	30/01/01
WO 9966040 A1	23/12/99	AU 4197499 A	05/01/00
		EP 1088068 A	04/04/01
WO 9717450 A2	15/05/97	AU 712650 B	11/11/99
		AU 1158297 A	29/05/97
		AU 7673796 A	29/05/97
		CA 2237039 A	15/05/97
		CA 2237120 A	15/05/97
		EP 0862640 A	09/09/98
		EP 0889966 A	13/01/99
		JP 2000500014 T	11/01/00
		JP 2000500015 T	11/01/00
		US 5716808 A	10/02/98
		US 5965389 A	12/10/99
		US 6001597 A	14/12/99
		US 6153424 A	28/11/00
		WO 9717451 A	15/05/97
		US 5888768 A	30/03/99
		US 5955349 A	21/09/99
WO 9632472 A1	17/10/96	AU 5333396 A	30/10/96
WO 9949022 A1	30/09/99	AU 3326699 A	18/10/99
		AU 3326799 A	18/10/99
		BR 9909006 A	28/11/00
		BR 9909010 A	28/11/00
		CN 1293542 T	02/05/01
		EP 1065941 A	10/01/01
		EP 1066373 A	10/01/01
		WO 9948380 A	30/09/99